

DISTINGUISHING KNOWLEDGE INTEGRATION GAINS IN EVOLUTION EDUCATION THROUGH CONCEPT MAP NETWORK ANALYSIS

*Beat A. Schwendimann, École Polytechnique Fédérale de Lausanne (EPFL),
Lausanne, Switzerland*

Email: beat.schwendimann@gmail.com, www.schwendimann.org

Abstract: Understanding how evolutionary changes happen requires connecting genotype and phenotype level concepts. This study uses a form of concept map, called Knowledge Integration Map (KIM), as a learning and assessment tool to support and assess connecting concepts within and across levels. KIMs visualize clusters and cross-links by dividing the drawing area into subject-specific spaces. A technology-enhanced curriculum on evolution has been developed and implemented in four high school science classrooms (number of students $n=94$). Pre- and posttests consisting of multiple choice items, short essays, and concept maps triangulated changes in students' integration of concepts related to evolution and their understanding of evolutionary changes of different organisms. Network analysis of concept maps identified changes in prominence of selected concepts and in cross-links between levels. Findings indicate that both low and high performing students made significant gains integrating concepts about the mechanisms of evolution. Results suggest that network analysis can be used to identify, compare, and track changes in concept maps over time.

1 Introduction

The theory of evolution is central to an integrated knowledge of biology (Thagard & Findlay, 2010) but teaching and learning evolution has been found difficult (Sinatra, Southerland, McConaughy, & Demastes, 2003). From a constructivist perspective, the theory of evolution is challenging to understand because learners can hold a wide variety of normative and non-normative concepts prior to formal instruction (Alters & Nelson, 2002). Many people show mixed reasoning using both normative and non-normative concepts of evolution (Evans 2005). From early childhood to adulthood, non-normative concepts are frequently used to explain why evolutionary changes happened (Shtulman 2006). Non-normative explanations often focus on the macro-level (phenotype) and fail connecting the underlying micro-level (genotype). Learners need scaffolded activities and tools to distinguish between non-normative and normative concepts and connect genotype and phenotype level concepts to explain how evolutionary changes happen. Connections between genotype and phenotype level concepts can be visualized through graphical representations, such as concept maps (Kinchin 2011). This study uses a biology-specific form of concept map, called Knowledge Integration Map (KIM) that distinguishes between genotype and phenotype level concepts and visualizes existing and missing connections between different levels (Schwendimann & Linn, 2015). This study explores KIMs as learning and assessment tools to support and distinguish learning gains through network analysis.

2 Methods

A technology-enhanced week-long unit on evolution, using human lactose intolerance as a case study, has been developed for this study and implemented in authentic science classrooms delivered through the online inquiry learning environment WISE (Linn, Clark, & Slotta, 2003). Guided by the knowledge integration (KI) pattern (Linn & Eylon, 2011), the WISE unit for this study was designed to elucidate the mechanisms (How do evolutionary changes happen? To whom do evolutionary changes happen?) of evolution. The knowledge integration pattern assumes that learners can hold multiple, often contradictory concepts of scientific phenomena and that learners can distinguish and make connections between new and existing concepts. Instruction guided by the KI pattern thus emphasizes eliciting existing alternative concepts (so that they can be inspected and considered), adding new normative concepts to the repertoire, helping students to build a set of criteria for distinguishing alternative concepts (for example through scaffolded inquiry activities), and encouraging students to sort out alternative concepts. According to the KI pattern view, learners need carefully designed instruction that makes new concepts accessible and guides them to revisit and distinguish alternative concepts, especially for complex and challenging topics such as evolution.

3 Curriculum Design

This study used human lactose intolerance as a case study that illustrates the mechanisms of evolutionary changes in a human context. Human lactose intolerance can illustrate the mechanisms of evolution [mechanisms] through changes in the production of the enzyme lactase. Building an integrated understanding of the theory of evolution requires learning how traits change (sources of variation), get selected (selection of variation), and are passed on

to offspring (inheritance) [mechanisms]. Understanding the mechanisms of evolutionary change is challenging because the modern theory of evolution consists of a complex network of concepts from different fields of science (for example genetics, cell biology, and population biology) and different levels (for example genotype and phenotype) (Duncan & Reiser, 2007). Research suggests that students have difficulties connecting concepts across different levels (Duncan & Tseng, 2011). Especially the genotype-phenotype duality is unique to biology and is fundamental to the understanding of heredity and development of organisms (Mayr, 1988). Genotype level concepts describe the genetic material and its variations over time due to random mutations and recombination. Phenotype level concepts refer to the expressed physical form of an organism that results from the genotype and environmental influences. The interplay between sources of random variation (genotype level) and selection (phenotype level) forms the basic understanding how changes get passed on to offspring (inheritance), which over time influences allele frequencies in the gene pool (Dawkins, 1976). Lacking integration of genotype and phenotype level concepts can be a major hindrance to understanding evolution (White, Heidemann, & Smith, 2013). A stronger integration of concepts related to evolution on the genotype-level (sources of variation) and phenotype-level (selection of variation) is expected to strengthen more frequent usage of normative concepts (such as ‘mutation’) and decrease the use of non-normative concepts.

3.1 Procedure

Human lactose tolerance served as the case study for a weeklong inquiry-driven evolution unit, called ‘Gene Pool Explorer’, delivered through the web-based inquiry-learning environment (WISE). The unit began with a short outline by the teacher followed by an introduction to the basic techniques of concept mapping. Students practiced generating concept maps in a task using an everyday example (“What does it take to have a pizza delivered?”).

Pretest: This study used a pretest-posttest design to measure students’ prior knowledge and track changes in students’ integration of evolution-relevant concepts. Both tests were delivered through the WISE platform. Students individually filled out the pretest and completed a concept mapping task (see pretest-posttest construction below) on the first day of the unit. Immediately after finishing the unit, students completed a posttest (identical to the pretest) and a posttest concept map activity.

Unit activities: The WISE ‘Gene Pool Explorer’ unit consisted of six activities. Each activity included texts with photos or videos followed by several automatically graded multiple choice items as well as short essay questions for which the teacher provided feedback by the next day. The teacher randomly grouped students in each class into dyads. Each dyad shared one computer and worked at its own pace.

This paper will focus on analysis of pretest-posttest differences. Analysis of embedded concept mapping activities will be reported separately.

3.2 Participants

The WISE unit ‘Gene Pool Explorer’ was implemented in four general biology classes in a high school with an ethnically and socio-economically diverse student population of 9th and 10th grade science students. The high school had an enrolment of around 2000 students and was located in the urban fringe of a large US city. School-wide, 12% of students received free or reduced price meals; four percent of students were classified as English Learners; 67% were White, 16% were Hispanic or Latino, 10% were Asian, and 3% were Black. All students were familiar with technology-enhanced learning environments and a few had used other WISE units before. Only students who completed pretest, posttest, the WISE unit, and the concept map activities were included in the analysis (total n=115; included in study n=94). No student opted out of participating in the study. Students spent six hours over one week (two separate 50-minutes and two double periods) to complete the unit. Throughout the unit, the teacher and a researcher provided support for technical and content-specific questions. Prior to implementation, the unit has been reviewed by teachers, students, and biology experts to evaluate the content, clarity, and alignment with the curriculum. The WISE unit has been revised based on their suggestions. The participating teacher was an experienced master teacher with nine years of teaching experience and previous experience in using WISE units. The teacher implemented the WISE unit ‘Gene Pool Explorer’ as an introduction to the subsequent topic of evolution after having completed several weeks of introduction to genetics.

3.3 Pretest-posttest construction

Knowledge integration provided the design rationale for both the curriculum unit and the assessment. Knowledge integration items were constructed to measure qualitative and quantitative changes in the connectedness between concepts (described as the level of knowledge integration). The pretest/posttest consisted of different instruments

to track changes in students' understanding of evolution: eleven constructed-response short-essays (explanation items) (Liu, Lee, & Linn, 2010) and a concept map generation item (see table 1). Eight of the explanation items were preceded by multiple-choice questions that included normative and common non-normative concepts.

Instrument	Measurement	Variable
Pre/posttest multiple-choice and short essay items	Integration of concepts and shift from non-normative towards normative concepts	Knowledge integration score
Knowledge Integration Map (KIM)	A) KIM indicator concepts indicate shifts in prominence of of normative concepts	Indicator concept prominence score
	B) KIM cross-links indicate integration of genotype/phenotype level concepts	Crosslink score

Table 1: Assessment instrument overview

3.4 Pretest-posttest concept map construction

For the pretest, posttest, and the embedded concept mapping activities, this study used a form of concept map, called Knowledge Integration Map (KIM), which highlights connecting concepts from different levels by dividing the drawing area in a genotype and a phenotype level into which students needed to sort out the provided concepts (Schwendimann 2015; Schwendimann & Linn, 2015) (see figure 1). Students were given eleven concepts belonging to either the genotype or phenotype level (*natural selection, DNA, genetic drift, population size, gene, environment, adaptation, mutation, genetic diversity in gene pool, increase in fitness of population, and allele*), which aligned with concepts in the pre/posttest items and the WISE unit. The provided selection of concepts aimed to model experts' understanding which concepts are most central to understanding evolution. KIMs were designed to support knowledge integration processes by eliciting existing concepts, adding new concepts to the map, distinguishing alternative concepts through revision tasks, connecting concepts within and across levels, and applying concepts by generating explanations of scientific phenomena. Students had to develop their own criteria how to identify important connections between concepts. Student dyads needed to negotiate how to categorize each concept (by placing it in the corresponding area (genotype or phenotype), which connections and arrows to create, and which link labels to add. Figure 1 illustrates the different elements of the KIM worksheet: (1), evolution-specific levels (genotype and phenotype) (2), instructions (3), given list of concepts (4), and a starter map (5). All KIM activities used the electronic concept mapping tool CmapTools (Cañas *et al*, 2004).

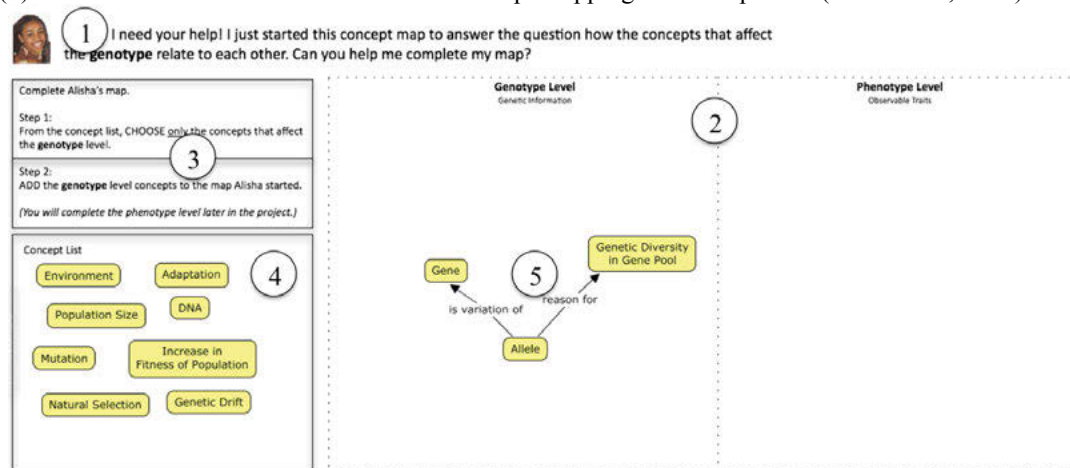


Figure 1: KIM worksheet

3.5 Pretest-posttest analysis

Student learning of concepts related to evolution was analyzed by comparing pretest and posttest scores. KI items distinguish the quality and quantity of connections that students construct among concepts. Pretest and posttest constructed-response items were scored according to a six-scale KI rubric (0-5) to measure changes in students' abilities to select normative over non-normative concepts and connect concepts (see table 2). A higher KI score indicates a higher number of connections between scientifically normative concepts and was interpreted as more integrated knowledge of scientific concepts. Explanations including a single concept or a mixture of correct and

incorrect elements were coded as ‘partial. ‘Basic’ explanations needed to connect two normative concepts with a correct link whereas ‘full’ explanations consisted of three or more correctly connected normative concepts. Each KI explanation item was independently scored by two raters: For the pretest, the ICC coefficient was 0.84, analysis of variance $F = 1.49$; and for the posttest, the ICC coefficient was 0.88, analysis of variance $F = 11.45$. For both tests, there was no significant difference between raters, and the level of concordance was high. The mean of the two ratings was used for the analysis and calculated as a total pre- and posttest score (maximum KI score $11 \times 5 = 55$). Each explanation item of the pre- and posttest was weighted equally. Changes in KI items from pretest to posttest were analyzed using paired t-tests with Holm-Bonferroni corrections to adjust for multiple comparisons (Holm 1979).

KI level	KI Score	Student Sample Answer
No Answer (blank)	0	None
Offtask	1	I don’t know.
Irrelevant/Incorrect	2	Finches develop new beaks to adapt to a new environment
Partial	3	Finches inherit traits from their parents.
Basic	4	Finches have differently shaped beaks that give them different chances to survive natural selection.
Complex	5	Natural selection causes those finches with helpful mutations to their beaks to be more genetically fit and adapt to the environment better. Therefore, the finches with the beaks adapted to their environment are more likely to reproduce and the trait gradually becomes dominant in the group.

Table 2: Knowledge integration sample rubric

For the post-hoc analysis, students were stratified into two performance groups according to their KI explanation item pretest scores (low and high prior knowledge of concepts related to evolution). To determine the effects of the unit on the two groups, a linear mixed effect analysis with robust standard errors was conducted using the mean posttest score as the dependent variable, the mean pretest score as the independent variable, prior knowledge (pretest performance dummy variable) as the grouping variable, and the classes as random variables. Mixed-effect linear models were chosen to take clustering due to nested data (students within classes) into account. Robust standard errors were calculated using the Huber-White sandwich method (Huber 1967; White 1980).

3.6 Pretest-posttest KIM analysis

To track shifts in students’ connections between concepts, each KIM proposition (relation between two concepts) was scored independently by two researchers using the six-scale KIM knowledge integration rubric (0-5) (propositional analysis). For the pretest KIM, the ICC coefficient was 0.96, analysis of variance $F = 4.92$; and for the posttest KIM, the ICC coefficient was 0.94, analysis of variance $F = 5.85$. For both tests, there was no significant difference between raters, and the level of concordance was high. The mean of the two ratings was used for the analysis (KIM overall KI score). Table 3 shows the KI rubric for the relationship between the concepts ‘mutation’ and ‘genetic diversity. All propositions were weighted equally. Analogous to the KI explanation item analysis, a linear mixed effect analysis with robust standard errors was conducted using the mean KIM overall KI posttest score as the dependent variable, the mean KIM overall KI pretest score as the independent variable, prior knowledge (pretest performance dummy variable) as the grouping variable, and the classes as random variables. Robust standard errors were calculated using the Huber-White sandwich method.

In addition to the KIM propositional analysis, two complementary analysis methods were used to track changes in students’ understanding. First, network analysis was used to identify changes in students’ perceived importance of the chosen concepts (prominence). One concept from each level was selected as the ‘indicator concept’: ‘mutation’ for the genotype level and ‘natural selection’ for the phenotype level. As students develop a more complex understanding, they might also identify certain concepts as more important and connect them more often. Increases in the frequency of usage of these selected concepts served as indicators for a more integrated and normative understanding. Changes for each indicator concept are reported using the variable ‘prominence score’, calculated by adding up the KI scores of all connections leading to or from each indicator concept (Schwendimann 2014).

KI Score	Link label quality	Link Arrow	Sample Propositions
0	None (No connection)	None (No connection)	
1	Wrong label	Wrong arrow direction	Genetic diversity includes mutation
2	a) No label b) Correct label c) Incorrect label	a) Only line b) Wrong arrow direction c) Correct arrow direction	a) Mutation -- genetic diversity b) Genetic diversity –contributes to > mutation c) Mutation – includes > genetic diversity
3	No label	Correct arrow direction	Mutation --> Genetic Diversity
4	Partially correct label	Correct arrow direction	Mutation – increases -> Genetic Diversity
5	Fully correct label	Correct arrow direction	Mutation – causes random changes in the genetic material which in turn increases -> Genetic Diversity

Table 3: KIM knowledge integration sample rubric

Second, links between concepts placed in the genotype and phenotype level (cross-links) are of particular interest as they can indicate “creative leaps” (Novak & Canas, 2006, p. 2) in understanding by the creators of a concept map and reasoning across ontologically different levels (Duncan & Reiser, 2007). This study interpreted cross-links as important indicators of students’ emerging knowledge integration as they link between genotype and phenotype level concepts. The sum of all cross-links are reported in the variable ‘KIM cross-link KI score’, calculated by adding up the KI scores of all cross-links. Changes in KIMs (overall proposition score, cross-link score, indicator concept prominence scores) from pretest to posttest were analyzed using paired t-tests with Holm-Bonferroni corrections to adjust for multiple comparisons. The same pretest performance groups (low/high) were used for the KIM explanation item analysis.

4 Results

Overall, students showed significant learning gains from pretest to posttest in the multiple choice and explanation items (see table 4).

KI		Pretest			Posttest			Paired t-test	Significance (with Holm-Bonferroni correction)	Effect size
	n	Min	Max	Mean (SD)	Min	Max	Mean (SD)	t (DF)	p-value (level)	Cohen’s d
KI multiple choice items	94	6	21	15.73 (2.68)	11	22	16.87 (2.38)	3.66 (93)	0.0004 (***)	0.45
KI explanation items	94	6	40	27.64 (5.90)	12	50	30.78 (5.62)	6.21 (93)	0.0000 (***)	0.63
KI normative evolution concepts variable	94	0	4	0.58 (0.92)	0	4	0.88 (1.06)	2.68 (93)	0.009 (**)	0.30

Table 4: Overview table of knowledge integration explanation items: All students (*p<0.05; **p<0.01; ***p<0.001)

Between-subjects analysis indicates that students in both strata (low/ high prior knowledge of evolution in the pretest) significantly gained in integrating normative evolution-related concepts from pretest to posttest (see figure 5). Findings suggest that especially low performing students gained considerably after using the WISE unit ‘Gene Pool Explorer’.

Knowledge integration items		Pretest			Posttest			Paired t-test	Significance (with Holm-Bonferroni correction)	Effect size
Condition	n	M i n	Max	Mean (SD)	Min	Max	Mean (SD)	t (DF)	p-value (level)	Cohen's d
Low prior knowledge	38	6	27	22.5 (5.34)	12	38	27.66 (5.34)	5.67 (37)	0.000 (***)	0.66
High prior knowledge	56	28	40	31.13 (3.01)	26	50	32.89 (4.79)	3.44 (55)	0.001 (**)	0.44
KIM overall KI score	91	0	58	28.63 (11.44)	0	89	39.56 (14.09)	7.50 (91)	0.000 (***)	0.85
KIM cross-link KI score	91	0	15	2.99 (3.57)	0	30	7.85 (5.47)	8.15 (91)	0.000 (***)	0.47

Table 5: Overview table of KI explanation items and KIMs by prior knowledge group (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

To determine the effects of low and high prior knowledge, a linear mixed effect model analysis indicates a regression coefficient of the mean KI posttest score variable of 0.61 (SE = 0.08), $p < 0.001$, which indicates that students with low and high prior knowledge improved in integrating concepts related to evolution from pretest to posttest after using the WISE unit '*Gene Pool Explorer*' (see figure 2).

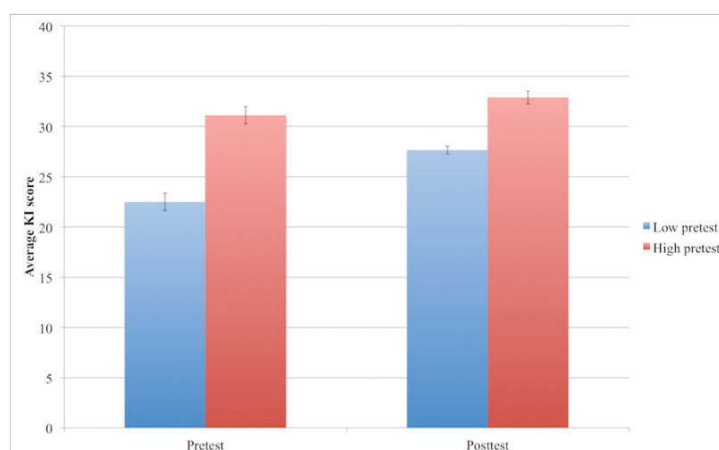


Figure 2: Pretest-posttest gains of low- and high-prior knowledge students

A composite variable for normative evolution concepts has been created from selected KI explanation items that focus on the mechanisms of evolution. The items were recoded to distinguish non-normative concepts (KI 0, 1, or 2), mixed concepts (KI 3), and normative concepts (KI 4 or 5). Paired t-tests with Holm-Bonferroni corrections suggest that students overall used normative evolution concepts more often than non-normative concepts in the posttest than in the pretest ($t(93) = 2.68$, $p < 0.01$).

KIM analysis: Overall, students improved the quantity and quality (KI scores) of connections between concepts in KIMs from pretest to posttest (see table 5). The prominence score for the two indicator concepts suggests that students made significant gains integrating these central normative evolution concepts (see table 6). Paired t-tests with Holm-Bonferroni corrections indicate that the prominence score for both KIM indicator concepts increased significantly from pretest to posttest: genotype level concept 'mutation' ($t(91) = 5.41$, $p = 0.000$ (two-tailed); pretest mean (SD) = 6.21 (4.51); posttest mean (SD) = 9.34 (5.21); Effect size (Cohen's d) = 0.31) and the phenotype level indicator concept 'natural selection' ($t(91) = 5.90$, $p = 0.000$ (two-tailed); pretest mean (SD) = 4.88 (3.43); posttest mean (SD) = 7.78 (4.70); Effect size (Cohen's d) = 0.33). Posttest results suggest that students placed the indicator concepts 'mutation' and 'natural selection' more often correctly in the

corresponding area, generated more connections to/from the indicator concepts, and that these connections were of higher quality (KI scores). These observations indicate that the normative concepts ‘mutation’ and ‘natural selection’ gained in explanatory strength (prominence) in students’ repertoire of concepts. Figure 3 shows a student example of changes of the prominence of indicator concepts in KIMs from pretest to posttest.

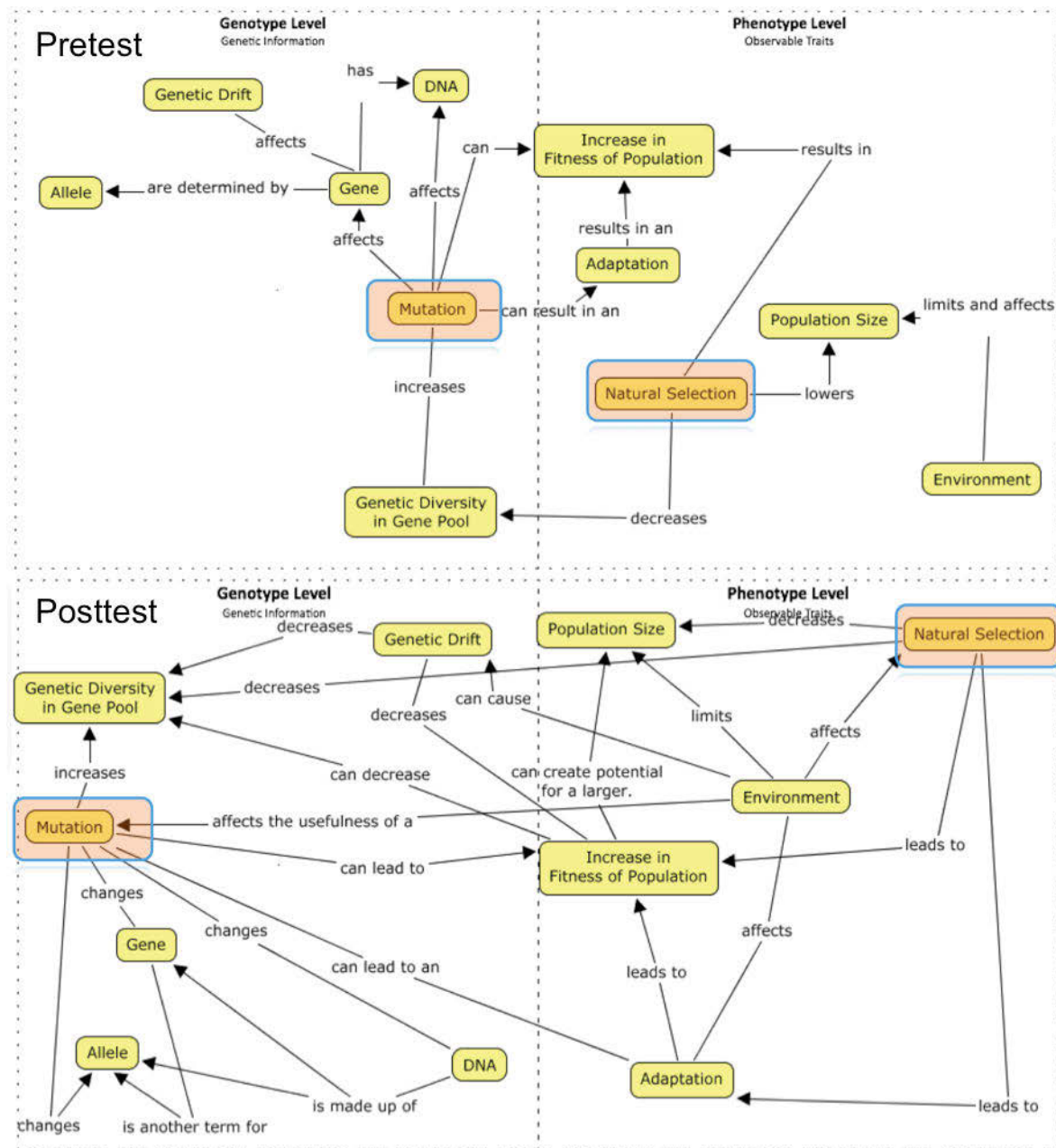


Figure 3: KIM student example (pretest and posttest)

To determine the effects of low and high prior knowledge, a linear mixed effect model analysis indicates a regression coefficient of the mean KIM posttest score variable of 0.51 (SE = 0.11), $p = 0.000$, which suggests that students with low and high prior knowledge generated better KIMs by connecting genotype and phenotype level concepts from pretest to posttest after using the WISE unit ‘*Gene Pool Explorer*’. These findings align with paired t-tests using Holm-Bonferroni corrections (see table 6). Both low and high prior knowledge students improved significantly in the prominence score for the indicator concepts ‘mutation’ and ‘natural selection’. Gains in prominence of the indicator concepts ‘mutation’ and ‘natural selection’ observed in KIMs align with increases in the KI composite variable for normative evolution concepts.

Variable		Pretest Pretest mean (SD)	Posttest Posttest mean (SD)	Paired t- test t (DF)	Significance Level (with Holm- Bonferroni correction) p-value (level)	Effect size Cohen's d
Overall	Mutation placement	0.61 (0.49)	0.84 (0.37)	4.25 (91)	0.0001 (**)	0.53
	Mutation number of links	1.76 (1.17)	2.53 (1.24)	5.20 (91)	0.000 (***)	0.64
	Mutation total KI score	6.21 (4.51)	9.34 (5.21)	5.41 (91)	0.000 (***)	0.64
By KI pretest knowledge	Mutation total KI score (Low prior knowledge)	5.73 (4.19)	8.30 (4.91)	2.74(36)	0.0095 (**)	0.56
	Mutation total KI score (High prior knowledge)	6.53 (4.72)	10.04 (5.33)	4.77 (54)	0.000 (***)	0.70

Table 6: Within-subjects and between-subjects statistics of changes in usage of indicator concept 'mutation'

KIM cross-link analysis: Results indicate that students significantly increased the number of cross-links between genotype and phenotype level concepts from pretest to posttest, (N=94): Pretest mean=1.03 (SD=1.15). Posttest mean=2.52 (SD=1.66). $t(93) = 7.49$, $p < .001$; effect size (Cohen's d) = 1.04 (which can be considered a large effect size) (see table 6). The average number of cross-links increased from 1 to 2.5. The number of KIMs with no cross-links was reduced from 43.6 % to 7.3 %. The KIM with the highest number of cross-links had four cross-links in the pretest and eight cross-links in the posttest. The KIM cross-link KI score indicates the quality of links between genotype and phenotype level concepts. The median (50th percentile) of the cross-link KI score increased from 2 to 7. The highest KI score of cross-links doubled from 15 to 30. This indicates that students did not only generate more cross-links but also propositions of higher quality. The increase of cross-links suggests that the WISE unit '*Gene Pool Explorer*' strengthened students' integration of genotype and phenotype level concepts. Mixed effect model analysis indicates a regression coefficient of the mean KIM cross-links score variable of 0.39 (SE = 0.16), $p = 0.01$, which suggests that students with low and high prior knowledge generated more cross-links connecting genotype and phenotype level concepts from pretest to posttest.

5 Conclusion and Discussion

Understanding the theory of evolution is crucial to understanding modern biology. To illustrate to mechanisms of evolution, this study distinguished between interrelated genotype (micro-level) and phenotype (macro-level) concepts, which were elicited in embedded KIMs and explored in scaffolded inquiry tasks. Findings from the pretest KI items and KIMs illustrate that students entered the unit with a rich repertoire of alternative concepts of evolutionary change, a fragmented understanding (indicated by the lack of cross-connections between genotype and phenotype level concepts). Initially, many students used non-normative concepts to explain evolutionary changes. Pretest-posttest gains, triangulated through different forms of assessment (multiple choice, explanation items, and KIMs), indicate that the WISE unit '*Gene Pool Explorer*' facilitated an improvement in knowledge integration and a shift towards normative concepts. Changes in KI explanation scores and KIM cross-link scores indicate that the WISE unit '*Gene Pool Explorer*' facilitated students' integration of genotype and phenotype level concepts of evolution. Findings suggest that a stronger integration of genotype and phenotype level concepts coincides with an increase in explanatory strength of normative concepts and a concomitant decrease of explanatory strength of non-normative concepts. Changes in explanatory strengths were identified through triangulating KI explanation items and KIM prominence scores. The analysis of posttest KIMs suggests that the normative concepts 'mutation' (genotype level) and 'natural selection' (phenotype level) gained in prominence in students' repertoire of concepts as they became more connected in students' maps. Increases in KIM cross-link

scores indicate stronger integration of concepts across the two levels. Post-hoc analysis of low and high prior knowledge groups suggests that the WISE unit '*Gene Pool Explorer*' can facilitate knowledge integration processes of all students, particularly learners with low initial understanding of evolution.

6 Limitations and Outlook

The WISE unit '*Gene Pool Explorer*' was implemented for a relatively short amount of time with a small sample size. A longer unit or a series of units with more participants could provide further insights into designing units that support knowledge integration processes of complex concepts in biology.

This study tracked changes in non-normative concepts. Explanations of evolutionary changes (in essays or concept map propositions) often use anthropomorphic language (Alters & Nelson, 2002). Even scientists and science educators frequently use anthropomorphic terms, such as 'need', 'desire', or 'intention' (Evans, Spiegel, Gram, & Diamond, 2009) to describe evolutionary change. Some argued that scientists might use teleological terms intentionally as shorthand or metaphors (Ariew 2003), while being aware of the underlying scientific processes. Legare and Evans found that need-based reasoning can provide a conceptual scaffold towards a scientific understanding (Legare, Lane, & Evans, 2013). However, using such accessible but possibly misleading language can also reinforce non-normative concepts of evolutionary theory (Jungwirth 1977). In both cases, learners need well-designed curricula and tools to distinguish normative and non-normative concepts, for example through scaffolded inquiry activities and knowledge visualization tools, such as KIMs (Schwendimann 2015) or MySystem (Linn & Chiu, 2011).

7 Implications

The WISE unit '*Gene Pool Explorer*' used a specific form of concept map, Knowledge Integration Maps (KIM), as summative and formative activities to elicit existing concepts and facilitate the generation of new connections. Like all forms of knowledge visualization, KIM tasks need to be preceded by an initial training phase to familiarize students with the basic techniques for generating and revising concept maps. KIMs aim to visualize and facilitate connections between genotype and phenotype level concepts. Collaboratively distinguishing concepts into different categories and highlighting cross-connections in KIMs can support knowledge integration processes of complex topics. When using KIMs, students might try to find the one "correct answer" for a KIM. Teachers should stress the point that each KIM is unique, and that there are many different possible solutions for a good KIM. This study used evolution-specific KIMs that distinguished between genotype and phenotype levels. By using different levels specific to other topics, KIMs could be adapted to support knowledge integration processes in other areas.

Acknowledgements. The research for this paper was supported by the National Science Foundation grant DRL-0334199 ("The Educational Accelerator: Technology Enhanced Learning in Science"). I thank my advisor Prof. Marcia C. Linn for her mentorship during the research for this paper and Prof. Pierre Dillenbourg for his support leading to the publication of this paper.

8 References

- Alters, B. J., & Nelson, C. E. (2002). Perspective: Teaching Evolution in Higher Education. *Evolution*, 56(10), 1891-1901.
- Ariew, A. (2003). Ernst Mayr's 'Ultimate/Proximate' Distinction Reconsidered and Reconstructed. *Biol Philos*, 18(4), 553-565.
- Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T., Lott, J, Carvajal, R. (2004). CmapTools: A Knowledge Modeling and Sharing Environment. In A. J. Cañas, J. D. Novak & F. M. González (Eds.), *Concept Maps: Theory, Methodology, Technology*. Proc. of the First Int. Conference on Concept Mapping (Vol. I, pp. 125-133). Pamplona, Spain: Universidad Pública de Navarra.
- Dawkins, R. (1976). *The Selfish Gene* (pp. xi, 224 p). New York: Oxford University Press.
- Duncan, R. G., & Reiser, B. J. (2007). Reasoning across Ontologically Distinct Levels: Students' Understandings of Molecular Genetics. *Journal of Research in Science Teaching*, 44(7), 938-959.
- Duncan, R. G., & Tseng, K. A. (2011). Designing Project-based Instruction to Foster Generative and Mechanistic Understandings in Genetics. *Science Education*, 95(1), 21-56.

- Evans, E. M. (2005). Teaching and Learning about Evolution. In J. Diamond (Ed.), *The Virus and the Whale: Explore Evolution in Creatures Small and Large* (pp. 25-40). NSTA Press Arlington, VA.
- Evans, E. M., Spiegel, A., Gram, W., & Diamond, J. (2009). Integrating Developmental and Free-choice Learning Frameworks to Investigate Conceptual Change in Visitor Understanding. *Ann Arbor*, 1001, 48190-5406.
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 65-70.
- Huber, P. J. (1967). The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 221-233).
- Kinchin, I. M. (2011). Visualising Knowledge Structures in Biology: Discipline, Curriculum and Student Understanding. *Journal of Biological Education*, 45(4), 183-189.
- Legare, C. H., Lane, J. D., & Evans, E. M. (2013). Anthropomorphizing Science: How does it Affect the Development of Evolutionary Concepts? *Merrill-Palmer Quarterly*, 59(2), 168-197.
- Linn, M. C., & Chiu, J. (2011). Combining Learning and Assessment to Improve Science Education. *Res Pract Assess*, 5, 4-13.
- Linn, M. C., & Eylon, B. -S. (2011). *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration* (p. 336). New York, NY: Routledge.
- Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE Design for Knowledge Integration. *Science Education*, 87(4), 517-538.
- Liu, O. L., Lee, H. S., & Linn, M. C. (2010). Multifaceted Assessment of Inquiry-based Science Learning. *Educational Assessment*, 15(2), 69-86.
- Mayr, E. (1988). *Towards a new Philosophy of Biology: Observations of an Evolutionist* (p. 564). Cambridge: Harvard University Press.
- Novak, J. D., & Cañas, A. J. (2006). *The Theory Underlying Concept maps and How to Construct Them*. IHMC.
- Schwendimann, B. A. (2014). Making Sense of Knowledge Integration Maps. In D. Ifenthaler & R. Hanewald (Eds.), *Digital Knowledge Maps in Education: Technology Enhanced Support for Teachers and Learners*. New York: Springer.
- Schwendimann, B. A. (2015). Concept Maps as Versatile Tools to Integrate Complex Concepts: From Kindergarten to Higher and Professional Education. *Knowledge Management & E-Learning: Special Issue on Novakian Concept Mapping in University and Professional Education*, 7(1), 73-99.
- Schwendimann, B. A., & Linn, M. C. (2015). Comparing two forms of Concept Map Critique Activities to Facilitate Knowledge Integration Processes in Evolution Education. *Journal of Research in Science Teaching*, 4.
- Shtulman, A. (2006). Qualitative differences between Naive and Scientific Theories of Evolution. *Cognitive Psychology*, 52(2), 170-194.
- Sinatra, G. M., Southerland, S. A., McConaughy, F., & Demastes, J. W. (2003). Intentions and Beliefs in Students' Understanding and Acceptance of Biological Evolution. *Journal of Research in Science Teaching*, 40(5), 510-528.
- Thagard, P., & Findlay, S. (2010). Getting to Darwin: Obstacles to accepting Evolution by Natural Selection. *Science & Education*, 19(6), 625-636.
- White, H. (1980). A Heteroskedasticity-consistent Covariance Matrix Estimator and a direct test for Heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817-838.
- White P. J. T., Heidemann M. K., & Smith J. J. (2013). A new Integrative Approach to Evolution Education. *BioScience*, 63: 586594.